JVI Journals.ASM.org

Conformational Plasticity of the 2A Proteinase from Enterovirus 71

Qixu Cai,[a] Muhammad Yameen,[*] Weihua Liu,[a] Zhenting Gao,[b] Yaozong Li,[b] Xuanjia Peng,[b] Yaxian Cai,[b] Caiming Wu,[a] Qian Zheng,[a] Jian Li,[b] Tianwei Lin[a]

State Key Laboratory of Cellular Stress Biology, School of Life Sciences, Xiamen University, Xiamen, China[a]; State Key Laboratory on Lead Compound Research, Wuxi AppTec Co., Ltd., Shanghai, China[b]

The 2A proteinase ($2A^{pro}$) is an enterovirally encoded cysteine protease that plays essential roles in both the processing of viral precursor polyprotein and the hijacking of host cell translation and other processes in the virus life cycle. Crystallographic studies of $2A^{pro}$ from enterovirus 71 (EV71) and its interaction with the substrate are reported here. EV71 $2A^{pro}$ was comprised of an N-terminal domain of a four-stranded antiparallel β sheet and a C-terminal domain of a six-stranded antiparallel β barrel with a tightly bound zinc atom. Unlike in other $2A^{pro}$ structures, there is an open cleft across the surface of the protein in an open conformation. As demonstrated by the crystallographic studies and modeling of the complex structure, the open cleft could be fitted with the substrate. On comparison $2A^{pro}$ of EV71 to those of the human rhinovirus 2 and coxsackievirus B4, the open conformation could be closed with a hinge motion in the bII2 and cII β strands. This was supported by molecular dynamic simulation. The structural variation among different $2A^{pro}$ structures indicates a conformational flexibility in the substrate-binding cleft. The open structure provides an accessible framework for the design and development of therapeutics against the viral target.

Enteroviruses are a family of single-stranded, positive-sense RNA viruses of *Picornaviridae* (1). Although most of the enterovirus-associated diseases are mild and asymptomatic, some members in the family can cause severe diseases and death, especially in the young and immunocompromised. Enteroviruses are the leading cause of aseptic meningitis, which in turn is the most common infection of the central nervous system (2). Enterovirus 71 (EV71) is an important pathogen, in addition to polioviruses, of the family. It is emerging as the most significant neurotropic enterovirus in some area of the world in outbreaks and epidemics of hand, foot, and mouth disease (HFMD) (3–5). This virus circulates in the United States, and 26% of the adults tested in a study had antibody (6). An epidemic of enterovirus 71 infection in 1998 in Taiwan resulted in over 400 cases of neurologic diseases and an estimated 1.5 million cases of HFMD and herpangina (7). Outbreaks of EV71-associated diseases have also been reported in the United States, Australia, Sweden, Japan, Bulgaria, Hungary, Hong Kong, Malaysia, Vietnam, and China (8–16). It has been associated with a variety of clinical diseases, including HFMD, herpangina, aseptic meningitis, encephalitis, and even fatal pulmonary edema or hemorrhage (17). Enterovirus can also cause severe chronic diseases, such as dermatomyositis, polymyositis, dilated cardiomyopathy, and diabetes mellitus (2).

Upon infection, a polyprotein is translated from the single open reading frame in the genome of an enterovirus and is processed into mature proteins by virally encoded 2A and 3C/3CD proteinases ($3C^{pro}$/$3CD^{pro}$) (1). The 2A proteinase ($2A^{pro}$) is an enzyme that cleaves at its own N terminus, at the junction between VP1 and 2A of the polyprotein. Besides its essential role in processing the viral proteins, $2A^{pro}$ shuts off the host's cap-dependent protein production by cleaving the elongation factors eIF4GI/II for the synthesis of viral proteins (18, 19). The $2A^{pro}$ of coxsackievirus B3 and other similar viruses in the enterovirus family cleaves dystrophin, a cytoskeletal protein, and is directly linked to dilated cardiomyopathy (20). $2A^{pro}$ from polioviruses also interferes with the nuclear traffic (21–23) and hijacks the splicing and transcription machinery (24). It is of great importance to understand the structural basis of $2A^{pro}$ function to assist in the design and development of therapeutics for the treatment of associated diseases. However, there is a gap in our knowledge of $2A^{pro}$, as its current structure reveals no space for substrate binding. Here we determined the crystal structures of EV71 $2A^{pro}$ and its complex with the substrate. The structure revealed an open substrate-binding cleft connecting from the S5 site to the active site and then extending to a possible S3′ site. Comparative studies with other $2A^{pro}$ proteins and molecular dynamic (MD) simulation indicated that the open conformation could be closed with a hinge motion in the bII2 and cII β strands. The open cleft provides an accessible template for docking chemical compounds and indicates a useful system for structure-based design and development of therapeutics to treat associated diseases.

MATERIALS AND METHODS

Protein expression and purification. The cDNA encoding EV71 $2A^{pro}$ was synthesized based on EV71 strain E2004104-TW-CDC (GenBank accession no. EF373576). It was amplified by PCR and inserted into the pGEX-4T-1 vector (GE Healthcare) to generate a plasmid containing the coding sequence of N-terminal glutathione *S*-transferase (GST) fusion $2A^{pro}$. A single mutation was made at the active site nucleophile (C110A). The cDNA encoding $2A^{pro}$ substrate peptide corresponding to the VP1-2A cleavage junction (ITTL*GKFG) was also inserted in the pGEX-4T-1 vector (GE Healthcare) to generate a plasmid containing the coding sequence of N-terminal GST fusion $2A^{pro}$ substrate. Following verification of its DNA sequence, *Escherichia coli* strain BL21(DE3) pLysS (Novagen) was transformed with the vectors pGEX-4T-1-$2A^{pro}$ and pGEX-

Received 26 December 2012 Accepted 12 April 2013
Published ahead of print 24 April 2013
Address correspondence to Tianwei Lin, twlin@xmu.edu.cn, or Jian Li, jian_li@wuxiapptec.com.
* Present address: Muhammad Yameen, Department of Chemistry, Government College University Faisalabad, Punjab, Pakistan.
Q.C., M.Y., and W.L. contributed equally to this article.

doi:10.1128/JVI.03541-12

4T-1-2Asub. The bacterial cultures were grown at 37°C to an optical density at 600 nm (OD_{600}) of 0.6, and 0.4 mM isopropyl-β-D-thiogalactoside (IPTG) was added to induce $2A^{pro}$ or substrate expression at 20°C for 12 h. Bacteria were subsequently harvested. For the purification of EV71 $2A^{pro}$, the bacteria were lysed by ultrasonication in lysis buffer (phosphate-buffered saline [PBS], pH 7.4). The lysate was centrifuged at 36,000 × *g* for 30 min to remove cell debris. The supernatant was loaded onto a glutathione-Sepharose column (GE Healthcare). After washing with PBS (pH 7.4) for several column volumes, the GST tag was removed with thrombin in PBS at 22°C overnight. A Superdex 75 16/60 column (GE Healthcare) was preequilibrated with a buffer containing 50 mM Tris-HCl (pH 8.0), and then the $2A^{pro}$ was loaded onto the column and the fractions containing $2A^{pro}$ were collected. The protein was concentrated to 10 mg/ml with a Centricon centrifugal filter unit (molecular weight cutoff [MWCO], 3,000) for crystallization. For the purification of EV71 $2A^{pro}$ substrate, the GST fusion $2A^{pro}$ substrate was loaded onto a glutathione-Sepharose column (GE Healthcare) and the GST tag was removed by thrombin. The $2A^{pro}$ substrate was collected and mixed with the purified $2A^{pro}$ at a molar ratio of 1:3 for 2 h. The complex was then concentrated to 10 mg/ml with a Centricon centrifugal filter unit (MWCO, 3,000) for crystallization.

For the enzymatic activity assay, the native sequence of EV71 $2A^{pro}$ was inserted in the pET-28a vector (Novagen) to generate a plasmid containing the coding sequence of the protein with a C-terminal 6×His tag. The vector was transformed into *Escherichia coli* strain BL21(DE3)/pLysS (Novagen) for protein expression with the same protocol as for pGEX-4T-1-$2A^{pro}$. The cells were harvested and lysed in PBS (pH 7.4), and the supernatant was loaded onto an Ni-nitrilotriacetic acid (NTA) column and eluted with 20 mM sodium phosphate (pH 7.4)–500 mM NaCl–150 mM imidazole. The protein solution was concentrated to 6 mg/ml and stored at −80°C for enzymatic activity assay.

Enzymatic activity assay. A fluorescent peptide with the sequence Dabcyl-KSRTAITTLGKFGQQSGE-Edans was employed as the substrate for the $2A^{pro}$ enzymatic and inhibition assays based on the fluorescence resonance energy transfer (FRET) effect. The sequence is derived from the VP1/2A junction of the polypeptide. The fluorescence was monitored at 500 nm with the excitation wavelength set at 340 nm using a fluorescence spectrophotometer (Cary Eclipse). The experiments were carried out in 20 mM Tris-HCl (pH 7.5) and 100 mM NaCl at 30°C. The measurements in the first 5 min were used to calculate the initial rates. The GraphPad Prism 5 program (GraphPad Software, Inc.) was used to calculate the K_m and 95% confidence intervals. The K_m was expressed as the best-fit value ± standard error of the mean (SEM).

Protein crystallization. Crystallization of EV71 apo-$2A^{pro}$ and complexes with substrate were carried out at 16 °C by hanging-drop vapor diffusion. The $2A^{pro}$ crystals were made with a well solution containing 0.1 M HEPES (pH 7.5), 20% 2-propanol, and 10% polyethylene glycol (PEG) 4000.

Substrate peptides were generated for crystallographic studies of complex structures. The sequence for the recombinant peptide was GSITTLG KFG. The sequences for the 10 synthetic peptides of 11 residues each were SRTAITTLGKF, SRTAITTMGKF, SRTAITTRGKF, SRTAITFLGKF, SRT AITFMGKF, SRTTITTLGKF, SRTFITTLGKF, SRTTITFLGKF, SRTTITT RGKF, and SRTFITTRGKF. The crystals of the $2A^{pro}$/substrate complex were obtained under conditions similar to those for apo-$2A^{pro}$.

X-ray data collection, processing, and structure determination. The crystals were transferred into a solution containing 0.1 M HEPES (pH 7.5), 20% 2-propanol, 10% PEG 4000, and 25% PEG 400 as a cryoprotectant prior to flash cooling to 100 K for data collection. X-ray diffraction data were collected using a Rigaku rotating-anode X-ray generator (λ = 1.5418 Å) equipped with a MAR Research 345-mm imaging plate and processed with the Automar package from Marresearch GmbH. Diffraction data were also obtained under synchrotron radiation, but the quality was not significantly improved and the resolution was not extended.

The structure of EV71 $2A^{pro}$ was determined by molecular replacement, employing the program Phaser (25) with the crystal structure of human rhinovirus 2 (HRV2) $2A^{pro}$ (Protein Data Bank [PDB] accession code 2HRV) as the search model. Manual model building and refinement were performed with Coot (26), CNS (27, 28), and Refmac5 (29, 30). To determine the $2A^{pro}$/substrate complex structure by molecular replacement, the refined $2A^{pro}$ structure was used as the initial phasing model. The substrate is located in the $2mF_o$-DF_c electron density map. The complex structure was refined with Refmac5. The statistics for data collection and refinement are given in Table 1. The presentations were made with the PyMOL Molecular Graphics System, version 1.4 (Schrödinger LLC).

TABLE 1 Data collection and refinement statistics

Parameter	Value[a] for: $2A^{pro}$	$2A^{pro}$/substrate complex
Data collection		
Space group	C2	C2
Cell dimensions		
a (Å)	86.613	86.566
b (Å)	44.265	43.991
c (Å)	52.036	51.764
β (°)	112.58	112.21
Resolution (Å)	50–1.90 (1.97–1.90)	50–1.66 (1.72–1.66)
R_{sym} or R_{merge}	0.0345 (0.2617)	0.0386 (0.4443)
$I/\sigma(I)$	11.9 (2.1)	11.5 (1.2)
Completeness (%)	92.4 (96.0)	93.3%(85.2)
Redundancy	1.93 (1.84)	3.66 (3.29)
Refinement		
Resolution (Å)	1.9	1.66
No. of reflections	12,668	18,996
R_{work}/R_{free}	0.194/0.240	0.190/0.226
No. of atoms		
Protein	1,069	1,098
Ligand/ion	1	1
Water	115	133
B factors (Å^2)		
Protein	29.3	26.3
Ligand/ion	38.0	27.0
Water	36.3	37.4
Ramachandran statistics (%)		
Most favored	93.38	97.08
Allowed	6.62	2.19
Outliers	0	0.73
RMSD		
Bond lengths (Å)	0.022	0.028
Bond angles (°)	2.02	2.10

[a] Values in parentheses are for the highest-resolution shell.

Modeling the interaction between $2A^{pro}$ and its substrate. The catalytic triads were superposed between the complex structure of EV71 $2A^{pro}$, with the partial substrate and the crystal structure of a $3C^{pro}$ from coxsackievirus A16 (CVA16), in complex with its substrate of 11 amino acid residues (P10 to P1) (Protein Data Bank accession code 3SJ9). Poly(Ala) of the P6-P3 moiety from the CVA16 $3C^{pro}$ complex structure was built onto the substrate sequence (P2-P2′) in EV71 $2A^{pro}$. After the residues were mutated to those of the native substrate sequence for EV71 $2A^{pro}$, an initial model for the P6-P2′ substrate was generated by Prime (Schrödinger) in the substrate-binding cleft of the EV71 $2A^{pro}$ structure. The model complex structure was further optimized using Prime (Schrödinger) MM-GBSA with default settings that freeze $2A^{pro}$ atoms but allow the substrate atoms to adjust during the energy minimization.

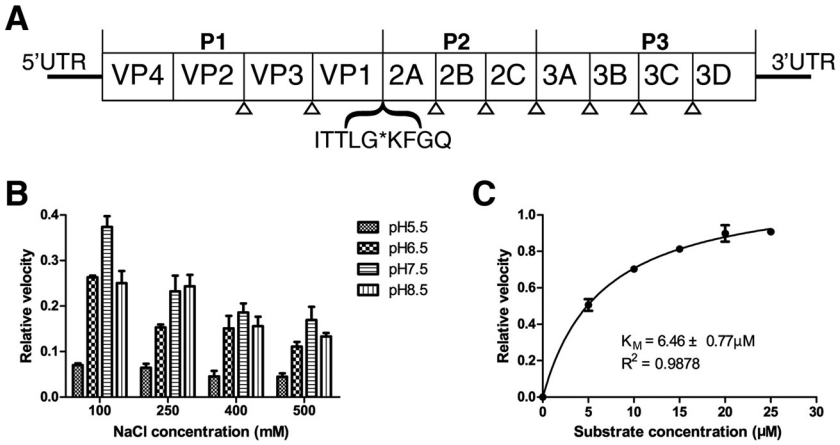

FIG 1 Enzymatic kinetic evaluation of EV71 2A^{pro}. (A) Diagrammatic representation of EV71 polyprotein. The substrate peptide sequence for 2A^{pro} was derived from the VP1/2A junction of the polypeptide. The cleavage site for 2A^{pro} is indicated by an asterisk, and the cleavage sites for EV71 3C^{pro} are indicated by triangles. UTR, untranslated region. (B) Optimization of the proteolytic reaction conditions. Two-dimensional grid searches were carried out by varying both pH and NaCl concentration. The optimal buffer condition is 20 mM Tris-HCl (pH 7.5) and 100 mM NaCl. (C) Measurement of K_m for EV71 2A^{pro} at 20mM Tris-HCl (pH 7.5) and 100 mM NaCl. Each data point was measured in triplicates. A K_m of 6.46 ± 0.77 μM was derived, and the 95% confidence interval for the K_m is between 4.82 and 8.10 μM.

Protein structure accession numbers. The coordinates of the crystal structures were deposited into the Protein Data Bank under accession numbers 4FVB and 4FVD.

RESULTS

Kinetic properties of EV71 2A^{pro}. The overexpressed and purified EV71 2A^{pro} was employed to investigate the kinetic properties of the enzyme by a method based on fluorescence resonance energy transfer (FRET) (31). Derived from the VP1/2A junction of the polypeptide (Fig. 1A), the substrate is a fluorescent peptide with the sequence Dabcyl-KSRTAITTLGKFGQQSGE-Edans. The optimal conditions for the catalytic activity were identified in a two-dimensional grid of varying pH and NaCl concentration. As shown in Fig. 1B, the best conditions for the proteolytic activity of EV71 2A^{pro} are 20 mM Tris-HCl (pH 7.5) and 100 mM NaCl at 30°C. A K_m of 6.46 ± 0.77 μM was derived, and the 95% confidence interval of the K_m is between 4.82 and 8.10 μM (Fig. 1C).

Structure of EV71 2A^{pro}. The sequence identity for the 2A^{pro} proteins of EV71 and coxsackievirus A16, two leading causative agents in HFMD, is 97%. That between EV71 and other coxsackieviruses A and B, including echoviruses, is about 75%. Polioviruses have a lower identity of about 60%. In contrast, the sequence identity with 2A^{pro} of human rhinoviruses, another group of picornaviruses encoding the enzyme, is below 40%. The sequences around the active site and zinc-binding residues are highly conserved among 2A^{pro} proteins from different viruses (Fig. 2A).

An EV71 2A^{pro} mutant in which the active site Cys110 was mutated to Ala was crystallized at pH 7.5, the optimal pH. The crystal structure was determined to 1.90-Å resolution by molecular replacement with 2A^{pro} from human rhinovirus 2 (HRV2) (32) as the initial phasing model (Table 1). The electron density is well defined for residues 7 to 144. The fold of EV71 2A^{pro} is similar to those of HRV2 2A^{pro} (32) and coxsackievirus B4 (CVB4) (33), with an N-terminal domain of a four-stranded antiparallel β sheet with a small α helix, which is a modified chymotrypsin fold (32), and a C-terminal domain of a six-stranded antiparallel β barrel with a tightly bound zinc atom (Fig. 2B).

There is a cleft across the surface of EV71 2A^{pro} with a constriction at one end where the active-site residues are located in the current structure (Fig. 3A). This is in contrast to the other structures of 2A^{pro}. The active site of EV71 2A^{pro} consists of a catalytic triad including His21 (His18 in HRV2 2A^{pro} and His21 in CBV4 2A^{pro}), Asp39 (Asp35 in HRV2 2A^{pro} and Asp39 in CBV4 2A^{pro}), and Cys110 (Cys106 in HRV2 2A^{pro} and Cys110 in CBV4 2A^{pro}) (Fig. 2). His21 functions as the general base, while Cys110 is the nucleophile which is replaced by Ala in this structure. The active-site residues can be superposed well between EV71 and HRV2 2A^{pro} (Fig. 2C).

In the Protein Data Bank, the CVB4 2A^{pro} structure is represented by 17 structures (PDB accession code 1Z8R). The root mean square deviation (RMSD) between Cα atoms in the 17 structures is between 1.6 to 3.8 Å. Most comparative studies for CVB4 2A^{pro} here were made with the first set of coordinates in 1Z8R. The RMSD between the Cα atoms of CVB4 and EV71 2A^{pro} is 9.0 Å if the structures are aligned by the LSQ routine in Coot (26). Better alignment could be obtained with the secondary-structure matching (SSM) routine (34) in Coot, and the corresponding core RMSD is 1.35 Å, which is the RMSD between the aligned Cα atoms. For comparison, the RMSD between the Cα atoms of HRV2 and EV71 2A^{pro} is 3.7 Å by LSQ, and the core RMSD is 1.14 Å by SSM. This is indicative that the core structures of CVB4 and EV71 2A^{pro} are similar. The variations in CVB4 2A^{pro} are a reflection of the dynamic nature of the enzyme (33). Since one member of the catalytic triad, Cys110, was in a region that was not well defined (33), no further comparison on the active-site residues from CVB4 2A^{pro} was made.

The cleft is long enough to accommodate the P5 to P1 residues of the substrate, using the nomenclature used by Schechter and Berger (35), and is connected to the S′ sites across the active site (Fig. 3A). If the active site is placed on the right, the south rim of the cleft is formed by the C-terminal domain with the active-site residue Cys110 under the rim in the narrow part of the cleft. While the N-terminal domain forms the northeast floor of the cleft along with the active residues of His21 and Asp39, the north rim is formed by the bII2-cII loop from the C-terminal domain (Fig. 2B

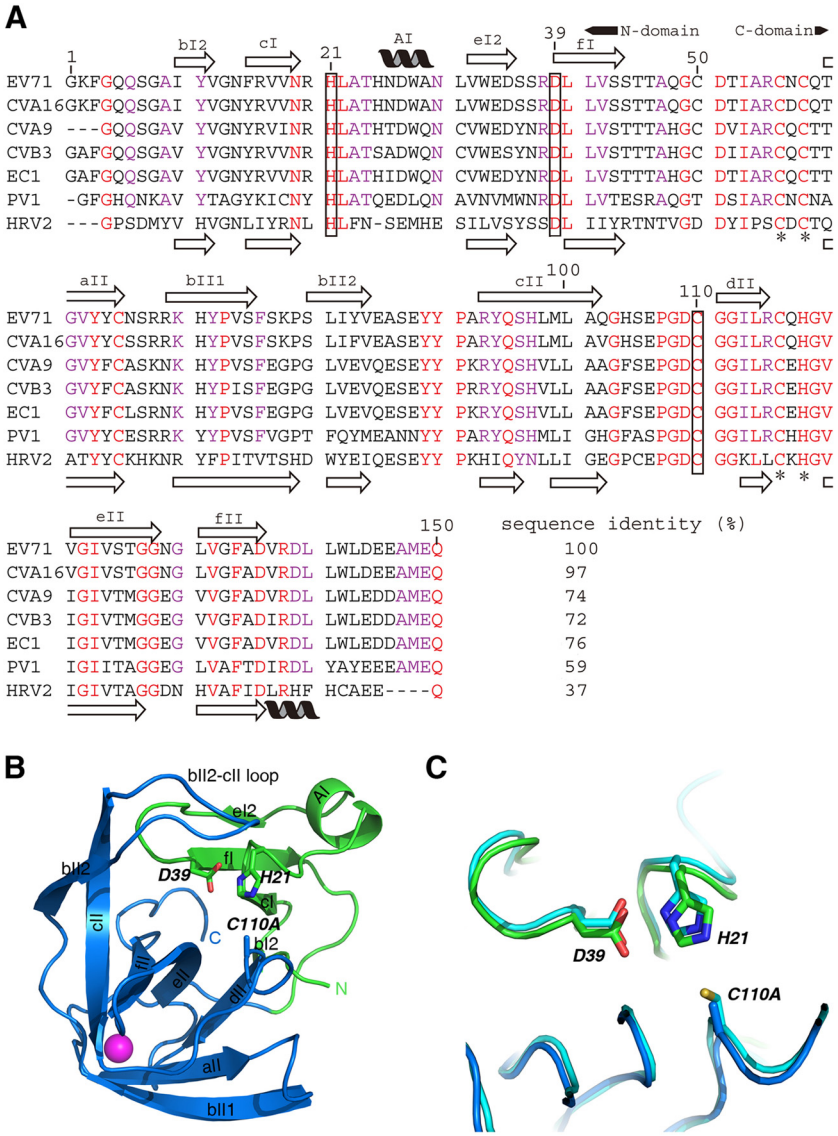

FIG 2 Sequence and structure of $2A^{pro}$. (A) Multiple-sequence alignment of $2A^{pro}$. All sequences were from the UniProt database. EV71, human enterovirus 71; CVA16, coxsackievirus A16; CVA9, coxsackievirus A9; CVB3, coxsackievirus B3; EC1, echovirus 1; PV1, poliovirus 1; HRV2, human rhinovirus 2. The residues identical among all viruses are in red, and those identical among enteroviruses are in purple. The active-site residues are in boxes. The residues in interaction with zinc are indicated by an asterisk underneath. The sequence identities are indicated. The secondary-structure assignments for EV71 $2A^{pro}$ are shown above the alignment, and those for HRV2 $2A^{pro}$ are below. (B) Ribbon diagram of EV71 $2A^{pro}$. The N-terminal domain is in green, and the C-terminal domain is in blue. The zinc ion is represented as a purple sphere. The active-site residues are shown. (C) Superposition of the active site residues between EV71 and HRV2 (cyan) $2A^{pro}$ proteins. The residue numbering is based on the sequence of EV71 $2A^{pro}$. Residue 110 of the catalytic triad in EV71 $2A^{pro}$ was mutated to Ala, while the corresponding residue in HRV2 $2A^{pro}$ is Cys.

and 3A). The S1 site is a shallow binding pocket with the oxyanion hole (32) underneath. It appears to be able to accommodate a side chain of moderate size. In the substrate-binding cleft, the S2 site is considerably larger and deeper under the north rim. This is not obvious in the other two $2A^{pro}$ structures, where the bII2-cII loop moves southward to compress the space (32, 33) (Fig. 3A and C). Although there is a clear opening in the substrate-binding cleft, it is still significantly narrower than the S1 and S2 sites of $3C^{pro}$, the other cysteine protease encoded in the EV71 genome (Fig. 3A and 4A). This is so because the bII2-cII loop is folded over the S2 site and active residues His21 and Asp39, while the β ribbon at the similar location in $3C^{pro}$ is not (36) (Fig. 3A and 4A and B). If the substrate-binding cleft in EV71 $2A^{pro}$ is described as a canyon, then the S1, S2, S1′, and active sites in $3C^{pro}$ form a basin. The distance from S1 to S2 across the canyon is about 10 Å (Fig. 3A), while the distance across the basin in the north-south direction is about 20 Å (Fig. 4A). This is indicative that the inhibitors targeting the active site of $3C^{pro}$ might not be suitable for $2A^{pro}$. The narrow cleft around the S1 and S2 sites, as well as across the active site, indicates a better chance of fitting noncovalent inhibitors in $2A^{pro}$. The S3, S4, and S5 sites are characterized by generally similar features that can be fitted with amino acid side chains along a groove. Another significant feature of the EV71 $2A^{pro}$ substrate binding is the uneven but accessible potholed contour surround-

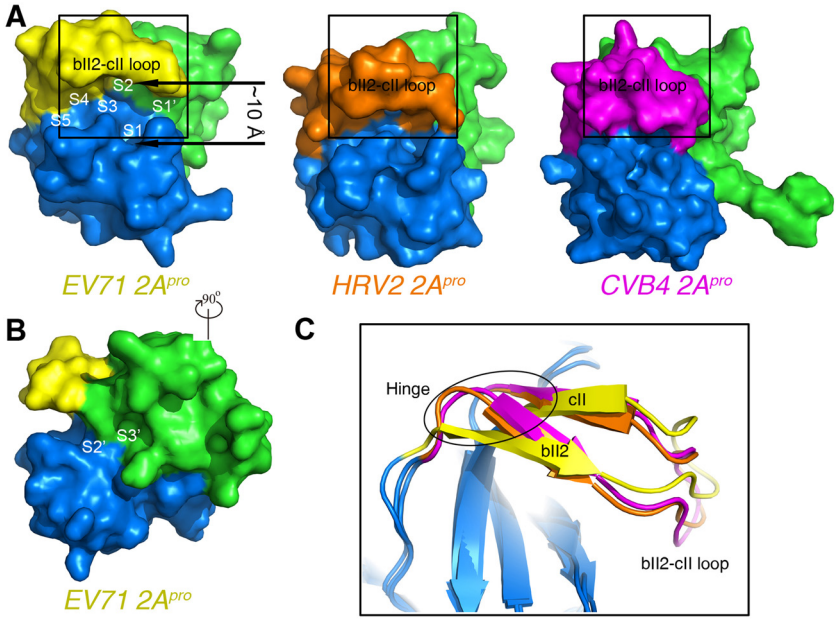

FIG 3 Comparison of 2A^{pro} proteins. (A) Surface rendering of 2A^{pro}. The N-terminal domain is in green, and the C-terminal domain, except the bII2-cII loop, is in blue. Left, EV71 2A^{pro}. There is an extensive cleft across the surface of the enzyme for substrate binding, with a constriction at the east end. The bII2-cII loop (yellow) forms the north part of the substrate-binding cleft, with catalytic residues under the constriction. The width of the cleft is about 10 Å. Middle, HRV2 2A^{pro}. The bII2-cII loop (orange) moves southward and the cleft becomes narrower and is closed over the constriction. The substrate peptides can no longer be fitted. Right, CVB4 2A^{pro}. The bII2-cII loop (magenta) blocks the access to the cleft from the exterior, and the substrate cannot be fitted. (B) A 90° view of EV71 2A^{pro} from panel A. A possible S3′ site is of a defined cavity and, together with S2′ site, well suited for fitting the small molecules. (C) Superposition of the bII2 and cII β strands and bII2-cII loop in a ribbon diagram. Parts of the bII2 and cII β strands in EV71 2A^{pro} change to coil-like structures in HRV2 and CVB4 2A^{pro} (circled), with color coding as in panel A. As a result, this part of the structure behaves like a hinge that moves the bII2-cII loop southward. The tip of bII2-cII loop in CVB4 2A^{pro} bends further downward.

ing the possible S3′ site (Fig. 3B), which is absent in 3C^{pro} (36) (Fig. 4). This difference is due to the absence of a similar aI-bI loop in 3C^{pro}, which covers the region (Fig. 4B).

An open conformation. Despite the overall similarity in fold, the north rims of the substrate-binding clefts of EV71, HRV2, and CVB4 2A^{pro} reveal striking differences in conformation (Fig. 3A and C). If the C-terminal domains are superposed, the bII2-cII loop of EV71 2A^{pro} in the north rim would have to slide southward to overlap the loops in HRV2 and CVB4 2A^{pro}. In addition, the tip of the loop would have to bend further downward to those of CVB4 2A^{pro}, where the substrate-binding cleft is all but inaccessible to substrate (Fig. 3A and C). The distance between the tips of the bII2-cII loops in EV71 and HRV2 2A^{pro} is about 3 Å. The distance between those in EV71 and CVB4 2A^{pro} is about 7 Å. The movement of the bII2-cII loop originates at a hinge motion of the bII2 and cII β strands, which unfold part of the fully extended β structures in EV71 2A^{pro} to make a kink in HRV2 and CVB4 2A^{pro} (Fig. 3A and C). The consequence of this closure of the cleft is that the substrate peptides can no longer be fitted into the current HRV2 and CVB4 2A^{pro} structures because of space con-

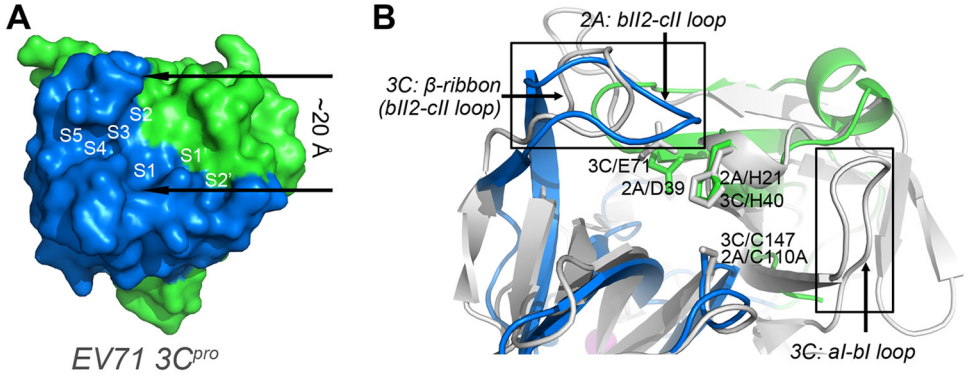

FIG 4 Comparison between EV71 2A^{pro} and 3C^{pro}. (A) Surface rendering of EV71 3C^{pro}. The N-terminal domain is in green, and the C-terminal domain is in blue. The active, S1, S2, and S1′ sites form a basin. The distance in the north-south crossing of the basin is about 20 Å. (B) Comparison of the active-site residues and substrate-binding clefts between 2A^{pro} and 3C^{pro} in ribbon diagrams. 3C^{pro} is in light gray. The bII2-cII loop in 2A^{pro} folds over the S2 site, while the aI-bI loop in 3C^{pro} covers the S′ sites.

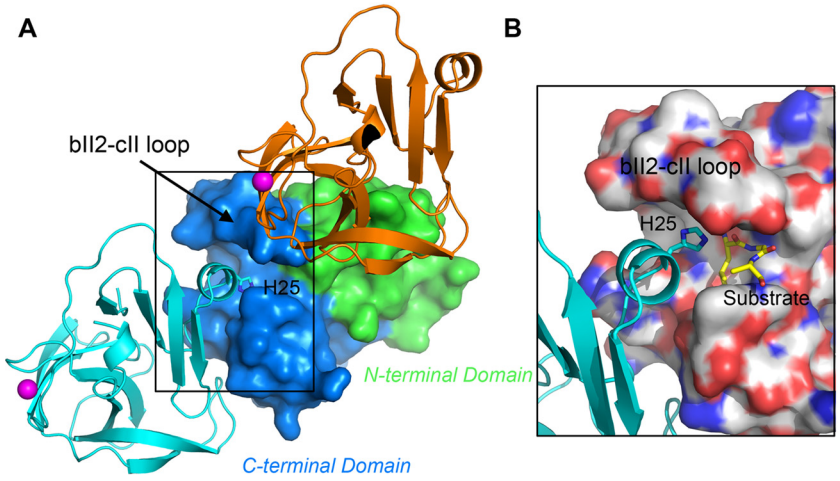

FIG 5 Intermolecular interactions in the EV71 2A^{pro} crystal. (A) There are 2 molecular contacts in the 2A^{pro} crystal. The neighboring molecule in orange ribbons interacts with the N-terminal domain, although it is close to the bII2-cII loop. The closest distance between the bII2-cII loop and the neighboring N-terminal domain is 5 Å. The molecule in cyan ribbons interacts with the south rim of the substrate-binding cleft, with residue His25 protruding into the binding cleft. (B) Binding of the substrate (in sticks) dislodges the neighboring His25, but there is no movement in either the south or north rim of the substrate-binding cleft. The area of this view is boxed in panel A.

straints (Fig. 3A). In light of the facts that the bII2-cII loop of HRV2 2A^{pro} is also variable between the two copies of the enzyme in the asymmetric unit of the crystal, the residues in this region have higher temperature factors (with an average of about 65 Å^2 for Cα atoms) and weaker electron density than other regions of HRV2 2A^{pro} structure (32), and CVB4 2A^{pro} is a substrate-free solution structure, it is conceivable that the bII2-cII loop of 2A^{pro} is a region susceptible to conformational change in response to the substrate binding, with the conformations found in the structures of HRV2 and CVB4 2A^{pro} for a substrate-free state and the conformation found in EV71 2A^{pro} representing an open conformation for binding the substrate, which is supported by the structural characterization of the 2A^{pro}/substrate complex (see below).

It is known that crystal packing has induced conformational changes in the structural elements directly involved in the intermolecular interaction in 3C^{pro} (36–39). The molecular contacts between EV71 2A^{pro} in the crystal were investigated. There are two types of intermolecular contacts in the EV71 2A^{pro} crystal (Fig. 5A). Neither of these contacts was directly associated with the bII2-cII loop. One of the contacts interacted with the N-terminal domain, although it was in the vicinity of the bII2-cII loop in the C-terminal domain (Fig. 5A). The closest distance between the bII2-cII loop and the neighboring N-terminal domain is 5 Å. The other contact involved the direct protrusion of the neighboring His25 into the substrate-binding cleft, but it interacted at the south rather than the north rim of the cleft (Fig. 5A). In addition, its dislodgement by the substrate did not alter the opening of the cleft (Fig. 5B). This indicated that the adoption of a different conformation by the bII2-cII loop was not an artifact of crystal packing. It might be argued that this interaction with the neighboring His25 could represent an initial event in the 2A^{pro} interaction with the substrate, however, in which the binding of a side chain from the substrate into the cleft would alter the dynamics of the structure to favor the open conformation and to make the cleft fully accessible to the substrate.

MD simulation. To evaluate the flexibility of the bII2-cII loop, an 8-nanosecond molecular dynamic (MD) simulation was carried out. EV71 2A^{pro} was placed in a periodic water box containing 150 mM NaCl to generate the MD trajectory. The RMSD for the Cα atoms, with respect to the first MD frame obtained after solvent equilibration, shows that the system became thermodynamically stable after approximately 3 ns in the MD simulation (Fig. 6A). The Cα distance between residues Glu88 and Pro107 was used as a measurement for the opening of the substrate-binding cleft. Glu88 is the residue at the tip of the bII2-cII loop. Pro107 is at the other side of the cleft in EV71 2A^{pro}. This distance varied dynamically between 8.5 and 16 Å in a periodic manner due to the movement of the bII2-cII loop (Fig. 6B).

An MD simulation with the same settings was also carried out for the HRV2 2A^{pro} (PDB accession code 2HRV). The result showed that HRV2 2A^{pro} is much more stable than EV71 2A^{pro} (Fig. 6C), although the movement between Cα of the equivalent Glu84 and Pro103 could be as large as 12 Å. The Cα distance between residues Glu88 and Pro107 in the EV71 2A^{pro} crystal structure is 11.1 Å, while the corresponding distance in HRV2 2A^{pro} crystal structure is 8.2 Å (Fig. 6D).

Interaction with the substrate. Eleven substrate peptides were employed for the crystallization of the complex. After many trials, a complex was obtained by cocrystallization with the Cys110Ala mutant and an octopeptide derived from the VP1-2A^{pro} junction of the precursor polypeptide. The structure was determined to 1.66-Å resolution in the same space group of the peptide-free enzyme. The EV71 2A^{pro} structure in the complex closely resembles the apo structure, with an RMSD of 0.24 Å. The average B factor was also similar for the bII2-cII loop at about 25 Å^2 for Cα atoms. The substrate interaction dislodged but did not completely displace the interaction from the neighboring His25 in the substrate-binding cleft (Fig. 5B). Probably as a consequence of this His25 interaction, which interferes with the substrate binding, the density was not fully connected (Fig. 7A), and only 4 residues of the substrate, with a sequence of TL/GK, were modeled in the electron density at the S2-S2′ sites (Fig. 7A). The P2 residue, a Thr, forms hydrogen bonds with residues Arg93 and Gln95 and is placed snugly in a pocket which can accommodate side chains

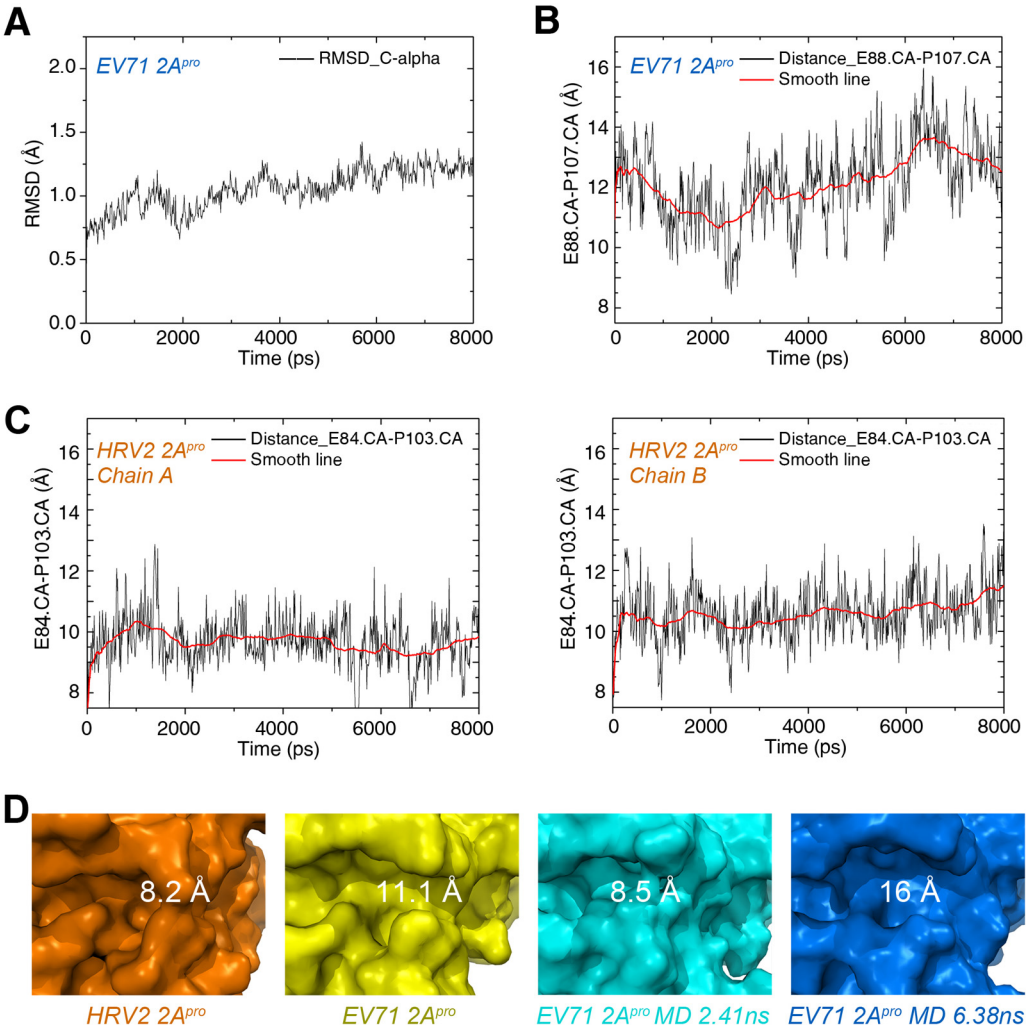

FIG 6 The cleft width varies during MD simulation. (A) Variation of RMSD for Cα atoms during the simulation. The system reaches equilibrium in 3 ns. (B) Distances between Cα atoms of Glu88 and Pro107 during the simulation. A smooth line is shown to highlight the trend. (C) Distance variations between Cα atoms of Glu84 and Pro103, two residues across the clefts in chain A (left) and chain B (right), of the HRV2 $2A^{pro}$ crystal structures during MD simulation. Both systems are more stable than EV71 $2A^{pro}$, although the change can be over 12 Å. (D) The distance between the Cα atoms of Glu84 and Pro103 across the cleft in HRV2 $2A^{pro}$ (orange) is 8.2 Å, and that of Glu88 and Pro107 in EV71 $2A^{pro}$ (yellow) is 11.1 Å, in the crystal structures. During the MD simulation, the shortest distance between the Cα atoms of Glu88 and Pro107 of EV71 $2A^{pro}$ is 8.5 Å at 2.41 ns (cyan), and the longest distance is 16 Å at 6.38 ns (blue).

larger than that of Thr. S1 is a shallow binding site, and its specificity toward Leu seems mainly due to space constraint around it. However, the space is open and can accommodate other residues in the upward direction, which probably explains the lack of specificity at S1 (40). A Gly residue at the P1′ site is probably due to the fact that the active site of the enzyme is highly constricted, and only the smallest residues can be accommodated. No interaction could be identified for P2′, probably because the electron density for its side chain is not well defined (Fig. 7A and B).

P6-P2′ substrate-binding model. Only a partial structure of the substrate was visualized in the complex. To get a better understanding of the $2A^{pro}$ interaction with the substrate, a model was generated for the full peptide binding at the cleft. Since both $2A^{pro}$ and $3C^{pro}$ are cysteine proteinases, it is conceivable that the structure of the P3-P6 moiety for $2A^{pro}$ could be referenced from the counterpart in $3C^{pro}$. The complex structures of EV71 $2A^{pro}$, with the tetrapeptide substrate and the complex structure between coxsackievirus A16 (CVA16) $3C^{pro}$ and its P1-P10 substrate (PDB accession code 3SJ9), were superposed on the catalytic triads and the P1-P2 moieties. The poly(Ala) derived from the P6-P3 moiety in $3C^{pro}$ was built onto the P2-P2′ substrate structure in the $2A^{pro}$ complex structure, and the initial model was made with the sequence TAITTL/GK, the sequence at the P1/P2 junction of the polyprotein. After energy minimization, the P6-P2′ substrate model of EV71 $2A^{pro}$ was generated (Fig. 8A). It is apparent that there are only small changes in the P1-P2 structure of the substrate in $2A^{pro}$. This suggests that the addition of other residues and the removal of His25 from the crystal packing would induce minimal perturbation to the peptide. The model substrate fit reasonably well into the substrate-binding cleft without any unfavorable geometry and clash between the enzyme and the substrate (Fig. 8A). The modeling of the substrate peptide onto the substrate-binding cleft in HRV2 $2A^{pro}$ would result in clashes between the peptide and the enzyme, especially around P2 (Fig. 8B). This is indicative

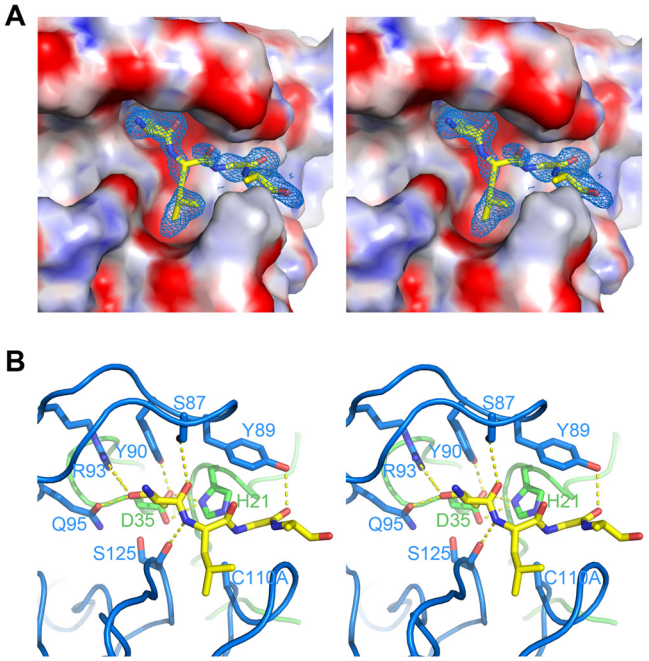

FIG 7 Complex structure of EV71 2A^{pro} and the substrate in stereo views. (A) Electrostatic surface potential of EV71 2A^{pro} contoured at between −10 kT/e (red) and +10 kT/e (blue) with the substrate and its electron density map ($2mF_o$-DF_c at 0.7 σ in blue chicken wire). (B) Interactions between EV71 2A^{pro} and the substrate. The hydrogen bonds or ionic interactions are indicated as dashed lines.

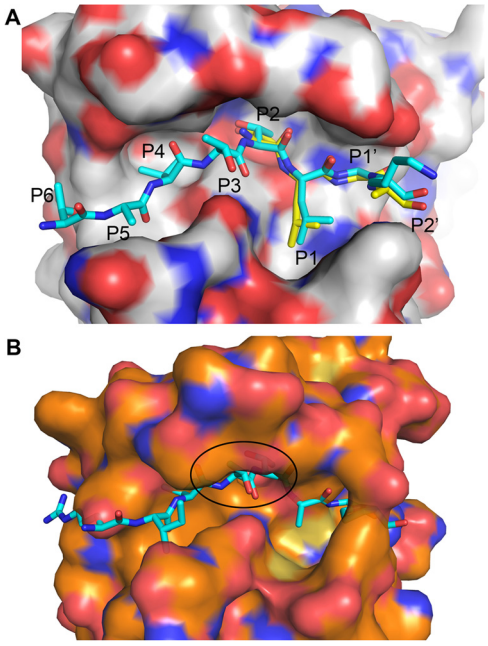

FIG 8 Model complex structures. (A) Modeling of the full substrate peptide in the binding cleft. Cα atoms in the model substrate are in cyan, and those in the crystal structure are in yellow. P2-P2′ of the model substrate matches closely to the crystal structure. The space constraint towards the exterior of the cleft is open for P1, which might explain the large variation among P1 residues. (B) Modeling of substrate peptide onto HRV2 2A^{pro}. There are steric clashes around the P2 site, highlighted by an oval. In order to bind the substrate in a fashion similar to that for EV71 2A^{pro}, the conformation of the bII2-cII loop would have to change.

that there would be a conformational change in the substrate-binding cleft of 2A^{pro} for the recognition of the substrate.

DISCUSSION

2A^{pro} plays multiple roles in enterovirus infection. On one hand, it is a vital component in the life cycle of the virus, as it cleaves its own N terminus during virus replication. On the other hand, it is essential for the viral usurpation or interference of host functions, such as the shutdown of cap-dependent translation of host proteins to facilitate the cap-independent translation of viral polyprotein (18, 19), the cleavage of dystrophin by that of coxsackievirus B3 (20), and the inhibition of nuclear traffic and splicing and transcription machineries by poliovirus 2A^{pro} (21–24). It is important to understand the structural basis of 2A^{pro} function and develop a framework for the design of therapeutics.

There are characteristic differences in the substrate-binding clefts between the two proteinases encoded in the enterovirus genome. This provides the structural basis for designing specific 2A^{pro} and 3C^{pro} inhibitors. The β ribbon in 3C^{pro} is peninsula-like with S2, S3, and S4 sites around it, but it does not fold over the S2 site (36). This results in a more exposed architecture around the active sites of S1 and S2. This more exposed architecture might be the reason that P2 is highly variable among 3C^{pro} substrates (41) and the reason for more tolerance to chemical groups taking different binding positions at S2 without diminishing the potency of a 3C^{pro} inhibitor (36). In contrast, the folding of the bII2-cII loop in 2A^{pro} over the S2 site and active-site residues makes the binding cleft more confined.

In this crystal structure, a neighboring His25 binds inside the substrate-binding cleft but not at the bII2-cII loop. It is not unreasonable to suggest that this interaction with a neighboring His25 might be analogous to an initial event in the substrate interaction with the enzyme. As the bII2-cII loop can vary among different conformations, as was demonstrated by MD simulation and the CVB4 2A^{pro} structure, an interaction with the side chain of the substrate peptide in the south rim of the substrate-binding cleft would tip the balance and facilitate the opening of the substrate-binding cleft, which lends support to the notion that the substrate binding induces the conformational change.

The more confined substrate-binding cleft compared to that in 3C^{pro} provides more constraint for fitting chemical compounds and a better chance to develop noncovalent inhibitors targeting EV71 2A^{pro}. Another distinct feature of the EV71 2A^{pro} structure is the potholed contour surrounding the possible S3′ site, which can be targeted for fitting compounds bridged across the active site. As there are no specifically designed 2A^{pro} activity inhibitors to date, the EV71 2A^{pro} structure provides a framework for the rational design and development of therapeutics to successfully target enterovirus-associated diseases.

ACKNOWLEDGMENTS

This research was supported by funding from the National Basic Research Program of China (973 program, 2012CB724500), from the Open Research Fund of the State Key Laboratory of Cellular Stress Biology, Xiamen University (SKLCSB2012KF003), and from the 111 Project of Education of China (no. B06016).

We gratefully acknowledge the crystallographic data collection at beamline BL17U1 at Shanghai Synchrotron Radiation Facility.

REFERENCES

1. **Racaniello VR.** 2006. *Picornavirdae*: the viruses and their replication, p 795. *In* Knipe DM, Howley PM (ed), Fields virology, 5th ed. Lippincott Williams & Wilkins Publishers, Hagerstown, MD.
2. **Pallansch M, Roos R.** 2006. Enteroviruses: polioviruses, coxsackieviruses, echoviruses, and newer enteroviruses, p 839. *In* Knipe DM, Howley PM (ed), Fields virology, 5th ed. Lippincott Williams & Wilkins Publishers, Hagerstown, MD.
3. **Bible JM, Pantelidis P, Chan PK, Tong CY.** 2007. Genetic evolution of enterovirus 71: epidemiological and pathological implications. Rev. Med. Virol. **17**:371–379.
4. **Solomon T, Lewthwaite P, Perera D, Cardosa MJ, McMinn P, Ooi MH.** 2010. Virology, epidemiology, pathogenesis, and control of enterovirus 71. Lancet Infect. Dis. **10**:778–790.
5. **Weng KF, Chen LL, Huang PN, Shih SR.** 2010. Neural pathogenesis of enterovirus 71 infection. Microbes Infect. **12**:505–510.
6. **Deibel R, Gross LL, Collins DN.** 1975. Isolation of a new enterovirus (38506). Proc. Soc. Exp. Biol. Med. **148**:203–207.
7. **Ho M, Chen ER, Hsu KH, Twu SJ, Chen KT, Tsai SF, Wang JR, Shih SR.** 1999. An epidemic of enterovirus 71 infection in Taiwan. Taiwan Enterovirus Epidemic Working Group. N. Engl. J. Med. **341**:929–935.
8. **Brown BA, Oberste MS, Alexander JP, Jr, Kennett ML, Pallansch MA.** 1999. Molecular epidemiology and evolution of enterovirus 71 strains isolated from 1970 to 1998. J. Virol. **73**:9969–9975.
9. **Chinese Center for Disease Control and Prevention and Office of the World Health Organization in China.** 2008, posting date. Report on the hand, foot and mouth disease outbreak in Fuyang City, Anhui Province and the prevention and control in China. http://www2.wpro.who.int/NR/rdonlyres/591D6A7B-FB15-4E94-A1E9-1D3381847D60/0/HFMDCCDC20080515ENG.pdf.
10. **Koroleva GA, Gracheva LA, Voroshilova MK.** 1978. Isolation of type 71 enterovirus from patients with a poliomyelitis-like disease during an outbreak in Bulgaria. Voprosy Virusolog **ii**:611–618.
11. **Melnick JL.** 1984. Enterovirus type 71 infections: a varied clinical pattern sometimes mimicking paralytic poliomyelitis. Rev. Infect. Dis. **6**(Suppl 2):S387–S390.
12. **Nagy G, Takatsy S, Kukan E, Mihaly I, Domok I.** 1982. Virological diagnosis of enterovirus type 71 infections: experiences gained during an epidemic of acute CNS diseases in Hungary in 1978. Arch. Virol. **71**:217–227.
13. **Samuda GM, Chang WK, Yeung CY, Tang PS.** 1987. Monoplegia caused by enterovirus 71: an outbreak in Hong Kong. Pediatr. Infect. Dis. J. **6**:206–208.
14. **Shimizu H, Utama A, Yoshii K, Yoshida H, Yoneyama T, Sinniah M, Yusof MA, Okuno Y, Okabe N, Shih SR, Chen HY, Wang GR, Kao CL, Chang KS, Miyamura T, Hagiwara A.** 1999. Enterovirus 71 from fatal and nonfatal cases of hand, foot and mouth disease epidemics in Malaysia, Japan and Taiwan in 1997-1998. Jpn. J. Infect. Dis. **52**:12–15.
15. **Shindarov LM, Chumakov MP, Voroshilova MK, Bojinov S, Vasilenko SM, Iordanov I, Kirov ID, Kamenov E, Leshchinskaya EV, Mitov G, Robinson IA, Sivchev S, Staikov S.** 1979. Epidemiological, clinical, and pathomorphological characteristics of epidemic poliomyelitis-like disease caused by enterovirus 71. J. Hyg. Epidemiol. Microbiol. Immunol. **23**:284–295.
16. **Yang F, Ren L, Xiong Z, Li J, Xiao Y, Zhao R, He Y, Bu G, Zhou S, Wang J, Qi J.** 2009. Enterovirus 71 outbreak in the People's Republic of China in 2008. J. Clin. Microbiol. **47**:2351–2352.
17. **Ho M.** 2000. Enterovirus 71: the virus, its infections and outbreaks. J. Microbiol. Immunol. Infect. **33**:205–216.
18. **Gradi A, Svitkin YV, Imataka H, Sonenberg N.** 1998. Proteolysis of human eukaryotic translation initiation factor eIF4GII, but not eIF4GI, coincides with the shutoff of host protein synthesis after poliovirus infection. Proc. Natl. Acad. Sci. U. S. A. **95**:11089–11094.
19. **Lloyd RE, Grubman MJ, Ehrenfeld E.** 1988. Relationship of p220 cleavage during picornavirus infection to 2A proteinase sequencing. J. Virol. **62**:4216–4223.
20. **Badorff C, Lee GH, Lamphear BJ, Martone ME, Campbell KP, Rhoads RE, Knowlton KU.** 1999. Enteroviral protease 2A cleaves dystrophin: evidence of cytoskeletal disruption in an acquired cardiomyopathy. Nat. Med. **5**:320–326.
21. **Gustin KE, Sarnow P.** 2001. Effects of poliovirus infection on nucleo-cytoplasmic trafficking and nuclear pore complex composition. EMBO J. **20**:240–249.
22. **Gustin KE, Sarnow P.** 2002. Inhibition of nuclear import and alteration of nuclear pore complex composition by rhinovirus. J. Virol. **76**:8787–8796.
23. **Park N, Katikaneni P, Skern T, Gustin KE.** 2008. Differential targeting of nuclear pore complex proteins in poliovirus-infected cells. J. Virol. **82**:1647–1655.
24. **Almstead LL, Sarnow P.** 2007. Inhibition of U snRNP assembly by a virus-encoded proteinase. Genes Dev. **21**:1086–1097.
25. **McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ.** 2007. Phaser crystallographic software. J. Appl. Crystallogr. **40**:658–674.
26. **Emsley P, Lohkamp B, Scott WG, Cowtan K.** 2010. Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. **66**:486–501.
27. **Brunger AT.** 2007. Version 1.2 of the Crystallography and NMR system. Nat. Protoc. **2**:2728–2733.
28. **Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL.** 1998. Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. D Biol. Crystallogr. **54**:905–921.
29. **Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F, Vagin AA.** 2011. REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr. D Biol. Crystallogr. **67**:355–367.
30. **Murshudov GN, Vagin AA, Dodson EJ.** 1997. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. D Biol. Crystallogr. **53**:240–255.
31. **Matayoshi ED, Wang GT, Krafft GA, Erickson J.** 1990. Novel fluorogenic substrates for assaying retroviral proteases by resonance energy transfer. Science **247**:954–958.
32. **Petersen JF, Cherney MM, Liebig HD, Skern T, Kuechler E, James MN.** 1999. The structure of the 2A proteinase from a common cold virus: a proteinase responsible for the shut-off of host-cell protein synthesis. EMBO J. **18**:5463–5475.
33. **Baxter NJ, Roetzer A, Liebig HD, Sedelnikova SE, Hounslow AM, Skern T, Waltho JP.** 2006. Structure and dynamics of coxsackievirus B4 2A proteinase, an enyzme involved in the etiology of heart disease. J. Virol. **80**:1451–1462.
34. **Krissinel E, Henrick K.** 2004. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. Acta Crystallogr. D Biol. Crystallogr. **60**:2256–2268.
35. **Schechter I, Berger A.** 1967. On the size of the active site in proteases. I. Papain. Biochem. Biophys. Res. Commun. **27**:157–162.
36. **Wu C, Cai Q, Chen C, Li N, Peng X, Cai Y, Yin K, Chen X, Wang X, Zhang R, Liu L, Chen S, Li J, Lin T.** 2013. Crystal structure of an enterovirus 3C proteinase (strain: E2004104-TW-CDC) and its complex with rupintrivir. Acta Crystallogr. D Biol. Crystallogr. **69**:866–871.
37. **Birtley JR, Knox SR, Jaulent AM, Brick P, Leatherbarrow RJ, Curry S.** 2005. Crystal structure of foot-and-mouth disease virus 3C protease. New insights into catalytic mechanism and cleavage specificity. J. Biol. Chem. **280**:11520–11527.
38. **Sweeney TR, Roque-Rosell N, Birtley JR, Leatherbarrow RJ, Curry S.** 2007. Structural and mutagenic analysis of foot-and-mouth disease virus 3C protease reveals the role of the beta-ribbon in proteolysis. J. Virol. **81**:115–124.
39. **Zunszain PA, Knox SR, Sweeney TR, Yang J, Roque-Rosell N, Belsham GJ, Leatherbarrow RJ, Curry S.** 2010. Insights into cleavage specificity from the crystal structure of foot-and-mouth disease virus 3C protease complexed with a peptide substrate. J. Mol. Biol. **395**:375–389.
40. **Sommergruber W, Ahorn H, Zophel A, Maurer-Fogy I, Fessl F, Schnorrenberg G, Liebig HD, Blaas D, Kuechler E, Skern T.** 1992. Cleavage specificity on synthetic peptide substrates of human rhinovirus 2 proteinase 2A. J. Biol. Chem. **267**:22639–22644.
41. **Lu G, Qi J, Chen Z, Xu X, Gao F, Lin D, Qian W, Liu H, Jiang H, Yan J, Gao GF.** 2011. Enterovirus 71 and coxsackievirus A16 3C proteases: binding to rupintrivir and their substrates and anti-hand, foot, and mouth disease virus drug design. J. Virol. **85**:10319–10331.